

Title: Method and Apparatus for Generating a Set of Filter Coefficients

Cross-references to Related Application

5 This application is related to the following applications:

1. United States Patent Application entitled, "Method and Apparatus for Generating a Set of Filter Coefficients for a Time Updated Adaptive Filter", filed on the same date as the instant application by Awad T. et al.
- 10 2. United States Patent Application entitled, "Method and Apparatus for Providing an Error Characterization Estimate of an Impulse Response Derived using Least Squares", filed on the same date as the instant application by Awad T. et al.
- 15 3. United States Patent Application entitled, "Method and Apparatus for Generating a Set of Filter Coefficients Providing Adaptive Noise Reduction", filed on the same date as the instant application by Awad T. et al.

The contents of the above noted documents are hereby incorporated by reference.

20 **Field of the Invention**

 The present invention relates generally to adaptive systems and, more particularly, to a method and apparatus for generating a set of filter coefficients. The method and apparatus are suitable for use in echo cancellation devices, equalizers, and in general, systems requiring time updated adaptive filtering.

30 **Background**

 Various adaptive filter structures have been developed for use in time updated adaptive systems to solve acoustical echo cancellation, channel equalization and other problems; examples of such structures include for example, transversal, multistage lattice, systolic array, and recursive implementations. Among these, transversal finite-impulse-

35

response (FIR) filters are often used, due to stability considerations, and to their versatility and ease of implementation. Many algorithms have also been developed to adapt these filters, including the least-mean-squares (LMS), recursive least-squares, sequential regression, and least-squares lattice algorithms.

5

A deficiency of many existing methods is that they generate new sets of filter coefficients iteratively, where the next set of filter coefficients depends on the previous set of filter coefficients. This is the case for the commonly used least-mean-square (LMS) algorithm. Occasionally, a newly generated set of filter coefficients is a worse
10 representation of the impulse response to be approximated than the previously generated set of filter coefficients. However, existing methods generally select the new set of filter coefficients and provide no suitable method for addressing this deficiency.

Consequently, there is a need in the industry for providing filter adaptation unit
15 suitable for producing a set of filter coefficients that alleviates, at least in part, the deficiencies of the prior art.

Summary of the Invention

20

In accordance with a broad aspect, the invention provides a method suitable for producing a set of filter coefficients indicative of an impulse response of a system in a given state. A sequence of samples of a first signal and a sequence of samples of a second signal are received. A first set of filter coefficients is generated on the basis of a
25 first subset of the sequence of samples of the first signal and a first subset of the sequence of samples of the second signal. The first set of filter coefficients is such that when an adaptive filter applies the first set of filter coefficients on the first signal, a first estimate of the certain component in the second signal is generated, the certain component being correlated to the first signal. A second set of filter coefficients is also generated on the
30 basis of a second subset of the sequence of samples of the first signal and a second subset of the sequence of samples of the second signal. At least some samples in the second subsets are subsequent to the samples in corresponding first subsets. The second set of filter coefficients is such that when an adaptive filter applies the second set of filter

coefficients on the first signal, a second estimate of the certain component in the second signal is generated, the certain component being correlated to the first signal. The first signal and the second signal are processed on the basis of the first set of filter coefficients to generate a first set of error characterization data elements and on the basis of the second set of filter coefficients to generate a second set of error characterization data elements. The first and second sets of filter coefficients and the second set of error characterization data elements are then processed to detect a change of state in the system. If a change of state is detected, the second set of filter coefficients is selected. If no change of state is detected, either one of the first set of filter coefficients and the second set of filter coefficients is selected, at least in part, on the basis of the first and second sets of error characterization data elements. A signal indicative of the selected set of filter coefficients is then released.

Advantageously, the invention allows detecting changes in the state of a system since the previously generating set of filter coefficients. In the non-limiting specific example of implementation of an echo canceller, this method allows the detection of echo point changes.

In a non-limiting example of implementation, the first subset of the sequence of samples of the first signal includes a plurality of successive samples, where the plurality of successive samples includes a beginning sample. Similarly, the second subset of the sequence of samples of the first signal includes a plurality of successive samples, where the plurality of successive samples includes a beginning sample. The detection of a change of state is prevented when the beginning sample of the first subset of the sequence of samples of the first signal is received subsequently to the beginning sample of the second subset of the sequence of samples of the first signal.

In other words, when the second subset of the sequence of samples of the first signal includes samples preceding the samples in the first subset of the sequence of samples of the first signal, then no change of state is detected. An advantage of this aspect is that it allows the newer samples to have a greater influence on the selection of the set of filter coefficients than the older samples. For example, consider the case where

the first set of filter coefficients is computed on the basis of samples of the first signal from time X to time Y and the second set of filter coefficients is computed on the basis of samples of the first signal from time A to time B. Since at least some samples in the second subset are subsequent to the samples in corresponding first subset, time B occurs later than time Y, or $B > Y$. In the case where A precedes X, or $A < X$, no change in the state of the system can be detected since any change of state should have already been taken into account by the first set of filter coefficients computed on the basis of the samples from time X to time Y. When A occurs subsequently to X, or $A > X$, a change in the state of the system may be detected.

10

In a non-limiting implementation, the first set of filter coefficients and the second set of filter coefficients are compared to derive a first distance measure data element. The first set of error characterization data elements is also compared to the second set of error characterization data elements to derive a second distance measure data element. A change of state is detected at least in part on the basis of the first distance measure data element and the second distance measure data element.

Continuing the non-limiting example of implementation, each error characterization data element in the first and second sets of error characterization data elements is associated to a respective frequency band selected from a set of frequency bands. The first set of filter coefficients is processed to derive a first set of spectral values, each spectral value corresponding to a respective frequency band selected from the set of frequency bands. The second set of filter coefficients is also processed to derive a second set of spectral values, each spectral value corresponding to a respective frequency band selected from the set of frequency bands. The first set of spectral values is compared to the second set of spectral values to derive a set of first distance measure data elements, where each of the first distance measure data elements is associated to a respective frequency band from the set of frequency bands. The first set of error characterization data elements and the second set of error characterization data elements are also compared to derive a set of second distance measure data element, each of second distance measure data elements being associated to a respective frequency band from the set of frequency bands. A change of state is detected, at least in part, on the basis of the

set of first distance measure data elements and the set of second distance measure data elements.

In a specific implementation, the set of frequency bands comprises one or more
5 frequency bands.

In a specific example of implementation, when no change of state is detected, a first energy data element is computed at least in part on the basis of the first set of error characterization data elements. The first energy data element is indicative of an error size
10 associated to the first set of filter coefficients. A second energy data element is also computed at least in part on the basis of the second set of error characterization data elements. The second energy data element is indicative of an error size associated to the second set of filter coefficients. Either one of the first set and the second set of filter coefficients is selected on the basis of the first energy data element and the second energy
15 data element.

Advantageously, the invention allows selecting between sets of filter coefficients on the basis of characteristics of the error function.

In a non-limiting example, the set of filter coefficients associated to the lowest
20 error size is selected.

In a specific example of implementation, the second set of filter coefficients is processed at least in part on the basis of a reference set of data elements to determine
25 whether the second set of filter coefficients is indicative of a substantially null transfer function.

In accordance with another broad aspect, the invention provides an apparatus for implementing the above-described method.

30

In accordance with yet another broad aspect, the invention provides a computer readable medium including a program element suitable for execution by a computing

apparatus for generating a set of filter coefficients in accordance with the above described method.

In accordance with another aspect, the invention provides an adaptive system.

5 The adaptive system includes first and second input for receiving a sequence of samples from a first signal and a sequence of samples of a second signal respectively. The second signal includes a certain component that is correlated to the first signal. The adaptive system also includes a filter adaptation unit and an adaptive filter. The filter adaptation unit has a coefficient generation unit for generating a first set of filter coefficients on the

10 basis of first subsets of the sequences of samples of the first signal and the second signal. The coefficient generation unit is also for generating a second set of filter coefficients on the basis of second subsets of the sequences of samples of the first signal and the second signal. At least some samples in the second subsets are subsequent to the samples in corresponding first subsets. An error characterization unit processes the first signal and

15 the second signal on the basis of the first set of filter coefficients to generate a first set of error characterization data elements associated to the first set of filter coefficients. The error characterization unit processes the first signal and the second signal on the basis of the second set of filter coefficients to generate a second set of error characterization data elements associated to the second set of filter coefficients. A selection unit processes the

20 first and second sets of filter coefficients and the second set of error characterization data elements to detect whether a change of state has occurred. When a change of state is detected, the second set of filter coefficients is selected. When no change of state is detected, either one of the first set of filter coefficients and the second set of filter coefficients is selected, at least in part, on the basis of the first and second sets of error

25 characterization data elements. An output releases a signal indicative of the set of filter coefficients selected by the selection unit. The adaptive filter applies a filtering operation to the first signal on the basis of the set of filter coefficients received from the filter adaptation unit and generates an estimate of the component in the second signal, the component being correlated to the first signal.

30

In accordance with another aspect, the invention provides an echo cancellor comprising the above-described adaptive system.

In accordance with another aspect, the invention provides a method and filter adaptation unit suitable for producing a set of filter coefficients. The filter adaptation unit has a first input for receiving a sequence of samples of a first signal and a second input for receiving a sequence of samples of a second signal. The second signal includes a certain component that is correlated to the first signal. A processing unit generates a plurality of sets of filter coefficients on the basis of the sequences of samples of the first signal and the second signal, each set of filter coefficients in the plurality of sets of filter coefficients being associated to a respective segment of the second signal. A selection unit selects at least one set of filter coefficients from the plurality of sets of filter coefficients at least in part on the basis of the plurality of sets of filter coefficients and the first signal. An output releases a signal indicative of the set of filter coefficients selected by the selection unit.

Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures.

Brief Description of the Drawings

Fig. 1 is a block diagram of a time adaptive system including a filter adaptation unit 100 in accordance with an embodiment of the present invention;

Fig. 2 is a block diagram of the filter adaptation unit of figure 1 in accordance with a specific example of implementation of the invention including a coefficient generation unit 200, an error characterization unit 202 and a coefficient set selection unit 204;

Fig. 3 is a functional block diagram of the coefficient generation unit 200 of figure 2 in accordance with an example of implementation of the invention;

Fig. 4 is a functional block diagram of the error characterization unit of figure 2 in accordance with an example of implementation of the invention;

Fig. 5 is a functional block diagram of a standard deviation computation unit suitable for use in the error characterization unit in accordance with a non-limiting example of implementation;

Fig. 6 is a functional block diagram of coefficient set selection unit 204 of figure 2 in accordance with an example of implementation of the invention;

10

Fig. 7a is a flow diagram of a process implemented by the coefficient set selection unit 204 in accordance with an example of implementation of the invention;

Fig. 7b is a flow diagram of a process implemented by the coefficient set selection unit 204 in accordance with an alternative example of implementation of the invention; and

Fig. 8 is a block diagram of an apparatus for generating a set of filter coefficients in accordance with a specific example of implementation of the invention.

20 Detailed Description

Figure 1 shows a time adaptive system 170 in accordance with an embodiment of the present invention. In one example of non-limiting implementation, the time adaptive system 170 is used to remove unwanted components of a return signal Z 102 from a forward signal Y 106. Typically, the return signal Z 102 passes through a system 150 and emerges in the form of a noise signal E 114 which corrupts the forward signal Y 106, resulting in a corrupted forward signal X 104. In a digital system, this corruption process may be modelled as a sample-by-sample addition performed by a conceptual adder 118. Thus, each sample of the corrupted forward signal X 104 is the sum of a component due to the (clean) forward signal Y 106 and another component due to the noise signal E 114 where the noise signal E 114 is correlated to the return signal Z 102.

A non-limiting use of the time adaptive system 170 is in the context of acoustical echo cancellation, for example, in a hands-free telephony system that includes a loudspeaker and a microphone. In this case, the forward signal Y 106 is a locally produced speech signal which is injected into the microphone (represented by conceptual adder 118), the return signal Z 102 is a remotely produced speech signal which is output by the loudspeaker, the system 150 is a room or car interior and the noise signal E 114 is a reverberated version of the return signal Z 102 which enters the same microphone used to pick up the forward signal Y 106. The corrupted forward signal X 104 is the sum of the signals input to the microphone, including the clean forward signal Y 106 as well as the reverberation represented by the noise signal E 114.

Another non-limiting use of the time adaptive system 170 is in the context of electric echo cancellation, for example, where the echo is caused by an analog/digital conversion on the transmission channel rather than by a signal reverberation in a closed space. In this case, the forward signal Y 106 is a locally produced speech signal which travels on the forward path of the communication channel, the return signal Z 102 is a remotely produced speech signal which travels on the return path of the communication channel, the system 150 is an analog/digital conversion unit and the noise signal E 114 is a reflected version of the return signal Z 102 which travels on the same forward path of the communication channel as the forward signal Y 106. The corrupted forward signal X 104 is the sum of the clean forward signal Y 106 as well as the noise signal E 114.

To cancel the correlative effect of the noise signal E 114 on the forward signal Y 106, there is provided a filter 110, suitably embodied as an adaptive digital filter. The filter 110 taps the return signal Z 102 (which feeds the system 150) and applies a filtering operation thereto. In one embodiment of the present invention, such a filtering operation can be performed by a finite impulse response (FIR) filter that produces a filtered signal F 112.

The filter 110 includes a plurality N of taps at which delayed versions of the return signal Z 102 are multiplied by respective filter coefficients, whose values are denoted h_j , $0 \leq j \leq N-1$. The N products are added together to produce the filter output at time T .

Simply stated, therefore, the filtered signal F 112 at a given instant in time is a weighted sum of the samples of the return signal Z 102 at various past instances.

The filter coefficients h_j are computed by a filter adaptation unit 100 configured to receive the return signal Z 102 and the corrupted forward signal X 104. The manner in which the filter adaptation unit 100 processes these signals to compute the filter coefficients h_j is described in greater detail herein below.

Mathematically, the filtered signal F 112 at the output of the filter 110 can be described by the following relationship:

Equation 1

$$f_t = \sum_{i=0}^{N-1} h_i z_{t-i}$$

where

t is the current sample time;

f_t is the value of the filtered signal F 112 at time t ;

h_j is the value of the j^{th} filter coefficient;

z_k is a sample of the return signal Z 102 at time k ; and

N is the length (i.e., the number of taps) of the filter 110.

For convenience, equation 1 may be represented in matrix form as follows:

Equation 2

$$f_t = \underline{h}^T \underline{z}_t$$

where the underscore indicates a vector or matrix, where the superscript “ T ” denotes the transpose (not to be confused with the sample time “ t ” used as a subscript) and where:

Equation 3

$$\underline{h} = \begin{bmatrix} h_0 \\ h_1 \\ \dots \\ h_{N-1} \end{bmatrix} \text{ and } \underline{z}_t = \begin{bmatrix} z_t \\ z_{t-1} \\ \dots \\ z_{t-(N-1)} \end{bmatrix}$$

The output of the filter 110, namely the filtered signal F 112, is subtracted on a sample-by-sample basis from the corrupted forward signal X 104 to yield an estimate, denoted Y* 108, of the clean forward signal Y 106. In a desirable situation, the filter coefficients h_j will be selected so as to cause the resultant signal Y* 108 to be “closer” to the clean forward signal Y 106 than corrupted forward signal X 104. For at least one optimal combination of filter coefficients, the resultant signal Y* 108 will be at its “closest” to the clean forward signal Y 106.

It is sometimes convenient to define “closeness” in terms of a least-squares problem. In particular, the optimal filter coefficients are obtained by solving an optimisation problem whose object it is to minimise, from among all possible combinations of filter coefficients h_j , the mean square difference between instantaneous values of the resultant signal Y* 108 and the clean forward signal Y 106. The actual value of the minimum mean-square error is typically not as important as the value of the optimal filter coefficients that allow such minimum to be reached.

A reasonable assumption is that noise signal E 114 adds energy to forward signal Y 106. Therefore an expression of the least square problem is to minimise the resultant signal Y* 108. Mathematically, the problem in question can be defined as follows:

Equation 4

$$\min_{\underline{h}} E[(y_k^*)^2],$$

where $E[\circ]_t$ denotes the expectation of the quantity “ \circ ” over a subset of time up until the current sample time t . For the purpose of this specific example, the expression $E[\circ]_t$, will denote the summation of the quantity “ \circ ” over a subset of time up until the current sample time t . Another commonly used notation is $\sum[\circ]_t$. Therefore, for the purpose of

5 this example the expressions $E[\circ]_t$ and $\sum[\circ]_t$ are used interchangeably.

Now, from Figure 1 it is noted that:

Equation 5

10

$$y_k^* = x_k - f_k = x_k - \underline{h}_k^T \underline{z}_k$$

and

Equation 6

$$x_k = y_k + e_k.$$

15

Therefore, the problem stated in Equation 4 becomes:

Equation 7

$$\min_{\underline{h}} E[(x_k - \underline{h}^T \underline{z}_k)^2]_t,$$

20

Expanding the term in square brackets, one obtains:

Equation 8

25

$$(x_k - \underline{h}^T \underline{z}_k)^2 = x_k^2 - 2x_k \underline{h}^T \underline{z}_k + (\underline{h}^T \underline{z}_k)^2.$$

Taking the expected value of both side of equation 8, one obtains:

Equation 9

$$E[(x_k - \underline{h}^T \underline{z}_k)^2]_t = E[x_k^2]_t - 2E[x_k \underline{h}^T \underline{z}_k]_t + E[\underline{h}^T \underline{z}_k \underline{z}_k^T \underline{h}]_t$$

Minimizing the above quantity leads to a solution for which the resultant signal Y* 108 will be at its minimum and likely at its “closest” to the clean forward signal Y 106. To minimize this quantity, one takes the derivative of the right-hand side of Equation 9 with respect to the filter coefficient vector \underline{h} and sets the result to zero, which yields the following:

Equation 10

$$\frac{d}{d\underline{h}} (E[x_k^2]_t - E[2x_k \underline{h}^T \underline{z}_k]_t + E[\underline{h}^T \underline{z}_k \underline{z}_k^T \underline{h}]_t) = -2E[x_k \underline{z}_k]_t + 2E[\underline{z}_k \underline{z}_k^T \underline{h}]_t = 0.$$

Thus, an “optimal” set of filter coefficients \underline{h}^*_j solves the set of equations defined by:

Equation 11

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{h}^* = E[x_k \underline{z}_k]_t.$$

15

It is noted that equation 11 expresses the filter coefficient optimisation problem in the form $A\underline{h} = B$, where $A = E[\underline{z}_k \underline{z}_k^T]_t$ and $B = E[x_k \underline{z}_k]_t$ and that the matrix A is symmetric and positive definite for a non-trivial signal Z 102. The usefulness of these facts will become apparent to a person of ordinary skill in the art upon consideration of later portions of this specification.

20

In a non-limiting example, matrix A is a matrix of dimension $N \times N$ and B is a vector of M elements where M is generally greater or equal to N ($M \geq N$). If $M = N$, a single vector \underline{h}^* can be computed from the equation 11. If $M > N$, then a plurality of vectors \underline{h}^* can be computed, where each vector \underline{h}^* is associated to a respective subset of N elements of vector B . Each subset of N elements is associated to a particular segment of signal X 104. The segments of signal X 104 associated to the subsets of N elements of vector B may be overlapping or mutually exclusive without detracting from the spirit of the invention. One set of filter coefficients from the plurality of vectors \underline{h}^* can then be

25

selected for release to filter 110. Specific examples of the manner in which the one set of filter coefficients may be selected from the plurality of vectors h^* will be described later on in the specification.

5 It can be noted that signal X = signal Y + signal E, and that the above equation 11 can be rewritten as follows:

Equation 12

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{h}^* = E[(y_k + e_i) \underline{z}_k]_t.$$

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{h}^* = E[e_k \underline{z}_k]_t + E[y_k \underline{z}_k]_t$$

10

In other words, we can separate the filter function defined by the set of filter coefficients h^*_j , $0 \leq j \leq N-1$ into two components. The first term on the right hand side of the equation $E[e_k \underline{z}_k]_t$ contributes to the desired filter behaviour since the filter 110 tries to obtain a filter such that signal F 112 equals signals E 114. Therefore, the second term

15 on the right hand side of the equation $E[y_k \underline{z}_k]_t$ contributes to the error behaviour of the filter 110. Therefore the error function can be expressed as follows:

Equation 13

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{error_function}^* = E[y_k \underline{z}_k]_t.$$

20

It will be readily observed that where signal Z 102 and signals Y 106 are perfectly uncorrelated, i.e. $E[y_k \underline{z}_k]_t = 0$ for all t, the error function is zero.

It is noted that since the properties of the signals Z 102 and X 104 change with

25 time, so too does the optimal combination of filter coefficients h^*_j , $0 \leq j \leq N-1$, which solves the above problem in Equation 11.

It is also appreciated that the state of system 150 and therefore the transfer function may also change over time and consequently, is a further factor in the optimal combination of filter coefficients h^*_j , $0 \leq j \leq N-1$.

5 Specific examples of the manner in which the changes in the state of system 150 and in the optimal combination of filter coefficients h^*_j , $0 \leq j \leq N-1$, are detected will now be described in greater detail with reference to Figure 2, which depicts the filter adaptation unit 100 in greater detail.

10 **Filter Adaptation unit 100**

The filter adaptation unit 100 includes a first input 252 for receiving a sequence of samples of a first signal Z 102, a second input 254 for receiving a sequence of samples of a second signal X 104, a coefficient generation unit 200, an error characterization unit
15 202, a coefficient set selection unit 204 and an output 256 for releasing an output signal indicative of a set of filter coefficients H 116. The filter adaptation unit 100 also includes a memory unit 240 for storing the last active set of filter coefficients selected by the coefficient set selection unit 204 along with a corresponding set of error characterization data elements.

20

Coefficient Generation Unit 200

The coefficient generation unit 200 receives the first signal Z 102 and the second signal X 104 from the first input 252 and the second input 254 respectively. The
25 coefficient generation unit 200 is operative to generate a set of filter coefficients H_{new} 206 at least in part on the basis of the first and second signals. In a specific example, the coefficient generation unit 200 applies a least squares method on the first and second signals to derive the set of filter coefficients H_{new} 206. The coefficient generation unit 200 generates a set of coefficients H_{new} 206 by solving equation 11 reproduced below:

Equation 11

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{h}^* = E[x_k \underline{z}_k]_t.$$

5

The coefficient generation unit 200 releases a new set of filter coefficients h_j^* , designated as *Hnew* 206 in figure 2.

Other specific examples of implementation of the coefficient generation unit 200 is described in United States Patent Application entitled, "Method and Apparatus for Generating a Set of Filter Coefficients for a Time Updated Adaptive Filter", filed on the same date as the instant application by Awad T. et al. The contents of this document are hereby incorporated by reference.

Figure 3 shows a specific non-limiting implementation of the coefficient generation unit 200. As shown, the coefficient generation unit 200 includes a context update module 300 and a filter coefficient computation unit 302.

The context update module 300 receives the sequence of samples of the first signal Z 102 and the sequence of samples of the second signal X 104. The context update module 300 generates and maintains contextual information of the first signal Z 102 and the second signal X 104. The context update module 300 maintains sufficient contextual information about signals Z 102 and X 104 to be able to derive $E[\underline{z}_k \underline{z}_k^T]_t$ and $E[x_k \underline{z}_k]_t$ for the current time t. For each new received sample of signals Z 102 and X 104 the contextual information is updated. This contextual information is then used by the filter coefficient computation unit 302 to generate the set of filter coefficients *Hnew* 206. The specific realization of the context update module 300 may vary from one implementation to the other without detracting from the spirit of the invention. For the purpose of this description, the contextual information comprises a first set of data elements and a second set of data elements, where the first set of data elements is indicative of the auto-correlation of signal Z 102 $E[\underline{z}_k \underline{z}_k^T]_t$. The second set of data elements is a set of cross-

correlation data elements $E[x_k z_k]_t$ of the first signal Z 102 with the second signal X 104. The first set of data elements and the second set of data elements may be periodically cleared to remove “old” samples from the context update module. In a non-limiting example used in the context of echo cancellation, the first set of data elements and the second set of data elements are cleared every minute.

The filter coefficient computation unit 302, makes use of the contextual information provided by the context update module 300 to generate the set of filter coefficients H_{new} 206. The frequency of the computation of the new set of filter coefficients H_{new} 206 may vary from one implementation to the other without detracting from the spirit of the invention. In a non-limiting example, a new set of filter coefficients H_{new} 206 is computed every L samples of signals Z 102, where L is ≥ 2 . The filter coefficient computation unit 302 solves equation 11 reproduced below:

Equation 11

$$E[z_k z_k^T]_t \underline{h}^* = E[x_k z_k]_t.$$

In a non-limiting example, the first set of data element can be represented by an $N \times N$ matrix “A” describing the expected auto-correlation of signal Z 102, $E[z_k z_k^T]_t$. Matrix “A” is symmetric and positive definite. The second set of data elements, indicative of the expected cross-correlation between signal Z 102 and signal X 104, can be represented by a vector “B” of M elements, $E[x_k z_k]_t$. Finally the set of filter coefficients can be represented by a third vector \underline{h}^* . The relationship between “A”, “B” and \underline{h}^* can be expressed by the following linear equation:

Equation 14

$$A \underline{h}^* = B$$

There are many known methods that can be used to solve a linear equation of the type described in the above equation and consequently these will not be described further

here. If $M = N$, a single vector \underline{h}^* can be computed from the above equation. Therefore, in the case where $N=M$, H_{new} 206 is one set of filter coefficients associated to a single segment of signal X 104. If $M > N$, a vector \underline{h}^* can be computed for each subset of N elements of vector "B", where each vector \underline{h}^* is associated to a particular segment of signal X 104 included in the subset of N elements of vector "B". It is to be noted that the segments of signal X 104 associated to the subsets of N elements of vector "B" may be overlapping or mutually exclusive without detracting from the spirit of the invention.

For a same $N \times N$ matrix A, when $M > N$, a plurality of subsets of N elements of vector "B" are used to generate a plurality of vectors \underline{h}^* , each vector \underline{h}^* being generated by the filter coefficient computation unit 302 by solving the above equation for a respective subset of N elements of vector B. Therefore, in the case where $N < M$, H_{new} 206 includes a plurality of sets of filter coefficients. H_{new} 206 can be expressed as a vector of sets of filter coefficients $H_{new}[i]$, for $i=0 \dots FC-1$, where FC is the number of sets of filter coefficients generated by solving equation 14, where each set of filter coefficients is associated to a respective segment of signal X 104. For example, the first set of filter coefficients $H_{new}[0]$, is generated on the basis of a first subset of the elements of vector B, namely $B[0 \dots N-1]$. The second set of filter coefficients $H_{new}[1]$, is generated on the basis of a second subset of the elements of vector B, namely $B[r \dots N-r-1]$ and so on and the last set of filter coefficients $H_{new}[FC-1]$ is generated on the basis of a last subset of the elements of vector $B[N-M \dots M-1]$.

Advantageously, it will be noted that although each set of filter coefficients $H_{new}[i]$ is generated on the basis of a respective subset of N elements of vector B, all sets of filter coefficients $H_{new}[i]$, for $i=0 \dots FC-1$ make use of a same $N \times N$ auto-correlation matrix A. This allows modelling a system 150 over a time delay of M samples of signal X 104 while solving equation 14 with a same matrix A of size $N \times N$. Specific examples of the manner in which a set of filter coefficients is selected from $H_{new}[i]$, for $i=0 \dots FC-1$, will be described in greater detail later on in the specification.

H_{new} 206 is generated on the basis of the contents of the context update module which includes data derived from a first subset of samples of signal Z 102 and a first

subset of samples of signal X 104. H_{new} 206 is associated to an ordering data element, which will be referred to as H_{new_age} , indicating the time of receipt of the first sample of signal Z 102 that was included in the computation of H_{new} 206. In a non-limiting example, when $M > N$, all sets of filter coefficients in $H_{new}[i]$, $i=0 \dots FC-1$, generated are
 5 assigned the same ordering data element, H_{new_age} .

H_{new} 206 is then released at the output 356 of the coefficient generation unit.

10 Error Characterization Unit 202

In accordance with a specific implementation, the error characterization unit 202 characterizes the error function associated with adaptive filter 170 on the basis of the knowledge of the amplitude of the first signal Z 102 and of an estimate of the amplitude
 15 of the forward signal Y 106.

Theoretical Explanation

As was described previously, the error function can be expressed by equation 13
 20 reproduced below:

Equation 13

$$E[\underline{z}_k \underline{z}_k^T]_t \underline{error_function}^* = E[y_k \underline{z}_k]_t.$$

25 In order to characterize the error function of the adaptive filter 170, a single tap filter is considered. In a single point tap system where $E[\underline{z}_i \underline{z}_i^T]_t$ has a single data element and $E[y_i \underline{z}_i]_t$ has a single data element, equation 13 can be written as follows:

Equation 15

$$30 \quad \sum_{i=0}^{t-1} z_i z_i \quad \underline{error_function} = \sum_{i=0}^{t-1} z_i y_i.$$

Solving for the error function at time t we obtain:

Equation 16

$$error_function(t) = \frac{\sum_{i=0}^{t-1} z_i y_i}{\sum_{i=0}^{t-1} z_i z_i}$$

5 where

t is the current sample time;

z_k is a sample of the return signal Z 102 at time k ;

and

y_k is a sample of the return signal Y 106 at time k ;

10

For the purpose of deriving a mathematical model to characterize the error function, an assumption is made that signal Z 102 and signal Y 106 are substantially independent of one another and are white. For the purpose of this specification, a signal S is white if $E(S_i S_j) \approx 0$ for $i \neq j$ and signals S and Q are independent if $E(S_i Q_j) \approx 0$ for all i, j .

15 The above assumptions allow considering that the error added by each sample pair is an independent variable which can be described by the following expression:

Equation 17

$$error_k = \frac{z_k y_k}{z_k z_k}$$

20 where z_k and y_k are the k th samples of signals Z 102 and Y 106 respectively and $error_k$ is the k th component of the error function due to the k th samples of signals Z 102 and Y 106. The error function can be considered as the sum of the errors added by the samples. In statistics, the above described error function can be considered to be a random variable. In order to characterize this random variable, the mean and the variance (or alternatively
25 the standard deviation) can be computed. Since signal Z 102 and signal Y 106 are assumed to be independent, the mean of this random variable is 0 and it will be shown below that the standard deviation can be given by:

Equation 18

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (z_i y_i)^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{t-1} z_i z_i \right)^2 - \left(\sum_{i=0}^{t-1} (z_i z_i)^2 \right) \right]^{\frac{1}{2}}}$$

5

Deriving the Standard Deviation Equation

The error inserted at each pair of samples $\{z_i, y_i\}$ can be represented mathematically by the following:

10

Equation 19

$$error = \frac{z_i y_i}{z_i z_i}$$

If the error components inserted at each pair of samples are equal to one another and are assigned equal weight, standard deviation of the error function after t samples can be expressed by the following expression:

15

Equation 20

$$\sigma_t = \frac{error}{\sqrt{t}} \quad \text{where } t \text{ is the number of samples}$$

and *error* is the error per sample

When each sample has an error that is different from the other and has a different weight, the standard deviation of the error function can be expressed as the division of two terms namely the average error over time and the number of samples conditioned by the weight. The average standard deviation of the error function can be expressed as follows:

20

Equation 21

$$\sigma_A = \frac{\left(\sum_{i=0}^{t-1} (error_{ii} * w_i)^2 \right)^{\frac{1}{2}}}{\left(\sum_{i=0}^{t-1} w_i^2 \right)^{\frac{1}{2}}}$$

- where w_i is a weight value associated to a given error component. The square root of the
 5 number of samples conditioned by the weight which corresponds to \sqrt{t} of Equation 20:

Equation 22

$$Sample\ number = \frac{\sum_{i=0}^{t-1} w_i}{\left(\sum_{i=0}^{t-1} w_i^2 \right)^{\frac{1}{2}}}$$

- 10 Therefore the standard deviation computation can be reduced to the following expression:

Equation 23

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (error_i * w_i)^2 \right)^{\frac{1}{2}}}{\sum_{i=0}^{t-1} w_i}$$

- 15 In a least squares context, the weight w_k of the error for each sample k is $z_k z_k$. Therefore, in the current specific example, the standard deviation of the error function can be expressed as follows:

Equation 24

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} \left(\frac{z_i y_i}{z_i z_t} * z_i z_t \right)^2 \right)^{\frac{1}{2}}}{\sum_{i=0}^{t-1} z_i z_t}$$

which can be reduced to the following:

5

Equation 25

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (z_i y_i)^2 \right)^{\frac{1}{2}}}{\sum_{i=0}^{t-1} z_i z_t}$$

10 In statistics, it is well known that when an unbiased estimator of the variance (or standard deviation) of a set of sample is to be obtained, the sample number is reduced by “1” to obtain an unbiased effective sample set. The effective sample set can be expressed by:

Equation 26

$$\text{Effective Sample number} = \left[\frac{\left(\sum_{i=0}^{t-1} w_i \right)^2}{\left(\sum_{i=0}^{t-1} w_i^2 \right)} - 1 \right]^{\frac{1}{2}}$$

15 Therefore the standard deviation computation can be reduced as follows:

Equation 27

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (error_i * w_i)^2 \right)^{\frac{1}{2}}}{\left(\sum_{i=0}^{t-1} w_i^2 \right)^{\frac{1}{2}}} * \frac{1}{\left[\frac{\left(\sum_{i=0}^{t-1} w_i \right)^2}{\left(\sum_{i=0}^{t-1} w_i^2 \right)} - 1 \right]^{\frac{1}{2}}} 1$$

5

Equation 28

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (error_i * w_i)^2 \right)^{\frac{1}{2}}}{\left(\sum_{i=0}^{t-1} w_i^2 \right)^{\frac{1}{2}}} * \frac{\left(\sum_{i=0}^{t-1} w_i^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{t-1} w_i \right)^2 - \left(\sum_{i=0}^{t-1} w_i^2 \right) \right]^{\frac{1}{2}}}$$

10

Equation 29

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (error_i * w_i)^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{t-1} w_i \right)^2 - \left(\sum_{i=0}^{t-1} w_i^2 \right) \right]^{\frac{1}{2}}}$$

In a least squares context, the weight w_k of the error for each sample k is $z_k z_k$. Therefore, in this second specific example, the standard deviation of the error function can be expressed as follows:

15

Equation 30

$$\sigma_t = \frac{\left(\sum_{i=0}^{t-1} (z_i y_i)^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{t-1} z_i z_i \right)^2 - \left(\sum_{i=0}^{t-1} (z_i z_i)^2 \right) \right]^{\frac{1}{2}}}$$

For the purpose of a specific implementation, equation 30 is used to characterize the
 5 standard deviation of the error function.

Adjustments for the Assumptions

As previously indicated, the above computations are based on the assumption that
 10 signals Z 102 and Y 106 are white and independent. The assumption that signal Z 102
 and signal Y 106 are independent is reasonable for many applications of adaptive
 filtering. It will be readily appreciated that when signal Z 102 and signal Y 106 are not
 exactly independent, the computations described in this specification may nevertheless be
 used with the knowledge that certain errors factors may be introduced by this
 15 approximation.

However, the assumption that signals Z 102 and Y 106 are white is not true in
 most applications. In order to solve this problem, signals Z 102 and Y 106 are divided
 spectrally into a set of frequency bands, where signal Z 102 and Y 106 can be considered
 20 to generally be substantially white within a given frequency band. In the non-limiting
 example of implementation of an echo cancellor, the signals Z 102 and Y 106 (assuming
 a sampling rate of 8000 samples/sec and therefore a frequency spectrum from 0 –
 4000Hz) are divided into 257 frequency bands of 15.625 Hz each. Using heuristics
 measurements, this width has been found to be narrow enough that voice is approximately
 25 a white signal across each of the 15.625 Hz bands. The width of the bands may vary from
 one application to another without detracting from the spirit of the invention. The
 “whiteness” of the signal is a subjective quality and depends on the nature of the signals
 being processed. The error function is then characterized for each frequency band

independently using the above described computation to estimate the mean (which is 0) and the standard deviation. For each frequency band, the standard deviation of the error function can be computed as follows:

Equation 31

$$\sigma_t[j] = \frac{\left(\sum_{i=0}^{t-1} (z_i[j] y_i[j])^2 \right)^{\frac{1}{2}}}{\sum_{i=0}^{t-1} z_i[j] z_i[j]}$$

where $z[j]$ and $y[j]$ is the amplitude of the component of signal Z 102 and signal Y 106 respectively in frequency band j and $\sigma_t[j]$ is the standard deviation of the error function in frequency band j at time t .

Another assumption in the above computations is that the amplitude (or energy) of signal Y 106 is known. However, signal Y 106 is unknown since, if signal Y 106 were known, the adaptive filter 170 would serve no practical purpose. The amplitude of signal Y 106 can be approximated by the amplitude of signal Y* 108. More specifically, in a least squares system, the forward signal Y 106 can be considered as made up of two (2) components, namely a first component Y_c which is correlated with signal Z 102 and a second component Y_u which is uncorrelated with signal Z 102. Because, by definition, Y_c and Y_u are uncorrelated, the energy of forward signal Y 106 is equal to the sum of the energies of Y_c and Y_u . Mathematically, this can be expressed as follows:

Equation 32

$$Y_{energy} = Y_c \text{ energy} + Y_u \text{ energy}$$

The filter 110 in combination with adder 180 will generally eliminate signal Y_c . Therefore, the energy of signal Y* 108 will be essentially equal to the Y_u energy which is less than or equal to the energy of signal Y 106. Therefore, since signal Y 106 is not available, the energy of signal Y* 108 is used as an approximation of the energy of signal

Y 106. For each frequency band, the standard deviation of the error function using Y* 108 can be computed as follows:

Equation 33

$$\sigma_t[j] = \frac{\left(\sum_{i=0}^{t-1} (z_i[j] y_i^*[j])^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{t-1} z_i[j] z_i[j] \right)^2 - \left(\sum_{i=0}^{t-1} (z_i[j] z_i[j])^2 \right) \right]^{\frac{1}{2}}}$$

5

Finally, although the above described standard deviation computations have been derived for an adaptive system having a single tap filter, similar derivations may be effected for a filter having N taps. In a practical application, for a filter having N taps, the standard deviation computation becomes:

10

Equation 34

$$\sigma_t[j] = \frac{\left(\sum_{i=0}^{\frac{t}{N}-1} (z_{N*i}[j] y_{N*i}^*[j])^2 \right)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{\frac{t}{N}-1} z_{N*i}[j] z_{N*i}[j] \right)^2 - \left(\sum_{i=0}^{\frac{t}{N}-1} (z_{N*i}[j] z_{N*i}[j])^2 \right) \right]^{\frac{1}{2}}}$$

15 In view of the above description, deriving a standard deviation computation for N > 1 will be readily apparent to the person skilled in the art and as such will not be described further.

Specific Implementation

20

As depicted in figure 2, the error characterization unit 202 receives the first signal Z 102, the second signal X 104 and the new set of filter coefficients H_{new} 206 from the

coefficient generation unit 200. The error characterization unit 202 is operative to generate at least in part on the basis of the first signal Z 102 and the second signal X 104 a set of error characterization data elements Herror 208 associated to the new set of filter coefficients *Hnew* 206. The error characterization unit 202 characterizes the error on a per-frequency band basis. Each error characterization data element in Herror 208 is a statistical estimate of the error function standard deviation for a respective frequency band.

Figure 4 shows a specific example of implementation of the error characterization unit 202 including a filter simulation unit 400, an adder unit 402, a first spectral calculator 406, a second spectral calculator 404 and a per-band standard deviation computation unit 408.

The filter simulation unit 400 is suitably embodied as an adaptive digital filter and simulates the processing of filter 110 shown in figure 1. The filter simulation unit 400 taps the return signal Z 102, and receives the new set of filter coefficients *Hnew* 206 from the coefficient generation unit 200. The filter simulation unit 400 applies a filtering operation corresponding to the filter coefficients *Hnew* 206 to the return signal Z 102 to produce filtered signal R 401. The manner in which the filtering operative is applied was described with regard to filter 110 in figure 1 and therefore will not be repeated here.

The output of the filter simulation unit 400, namely the filtered signal R 401, is subtracted by unit 402 on a sample-by-sample basis from the corrupted forward signal X 104 to yield a signal denoted W 470. Signal W 470 is an estimate of signal Y 106 (figure 1) generated on the basis of the set of filter coefficients *Hnew* 206.

First spectral calculator 406 taps the first signal Z 102 and divides the signal into a set of frequency bands. In a non-limiting example, the spectral calculator processes a set of samples of signal Z 102 from which the set of filter coefficients *Hnew* 206 was generated, where the first sample of the set of samples was taken at normalised time $t=1$. The spectral calculator 406 applies a set of Fast Fourier Transform (FFT) of length $(K-1)*2$, each Fast Fourier Transform (FFT) being applied to N of the samples of taps of the

adaptive filter 170. The computation of an FFT is well known in the art to which this invention pertains and as such will not be described further herein. For a given time t , the above calculation results into t/N sets of K spectral values of signal Z 102, each spectral value being associated to a respective frequency band from a set of K frequency bands.

- 5 In a non-limiting example used in echo cancellation, $K = 257$ is used to divide the frequency spectrum of signal Z 102 into 257 frequency bands. If the frequency spectrum goes from 0Hz to 4000Hz (assuming a sampling rate of 8000Hz), then there will be frequency bands centered at 0Hz, 15.625Hz, 15.625*2 Hz, 15.625*3 Hz, [. . .] and 4000Hz.

10

Mathematically, this can be expressed as follows:

Equation 35

$$Z_{SPECTRA} = FFT(z(t)) = \begin{bmatrix} Z_{SPECTRA}(0) \\ Z_{SPECTRA}(1) \\ Z_{SPECTRA}(2) \\ \dots \\ Z_{SPECTRA}(K-2) \\ Z_{SPECTRA}(K-1) \end{bmatrix}_0 \begin{bmatrix} Z_{SPECTRA}(0) \\ Z_{SPECTRA}(1) \\ Z_{SPECTRA}(2) \\ \dots \\ Z_{SPECTRA}(K-2) \\ Z_{SPECTRA}(K-1) \end{bmatrix}_1 \dots \begin{bmatrix} Z_{SPECTRA}(0) \\ Z_{SPECTRA}(1) \\ Z_{SPECTRA}(2) \\ \dots \\ Z_{SPECTRA}(K-2) \\ Z_{SPECTRA}(K-1) \end{bmatrix}_{\frac{t}{N}-1}$$

15

where

$$Z_{j,SPECTRA} = \begin{bmatrix} Z_{SPECTRA}(0) \\ Z_{SPECTRA}(1) \\ Z_{SPECTRA}(2) \\ \dots \\ Z_{SPECTRA}(K-2) \\ Z_{SPECTRA}(K-1) \end{bmatrix}_j$$

where $Z_{SPECTRA}$ 410 is a data structure of t/N vectors each of size K , each vector being indicative of a spectral representation of N samples of signal $z(t)$ and $Z_{SPECTRA}(j)$ is the spectral value of signal Z 102 associated to frequency band j . $Z_{SPECTRA}$ 410 is released

20 by first spectral calculator 406.

Second spectral calculator 404 taps the signal W 470 and divides the signal into a set of K frequency bands. In a non-limiting example, the second spectral calculator 404 processes a set of samples of signal W 470 corresponding to the set of samples of Z 102 processed by first spectral calculator 406, where the first sample of the set of samples of signal W 470 was taken at normalised time $t=1$. The first spectral calculator 406 applies a set of Fast Fourier Transform (FFT) of length $(K-1)*2$, each Fast Fourier Transform (FFT) being applied to N of the samples of signal W 470 where N is the number of taps of the adaptive filter 170. The computation of an FFT is well known in the art to which this invention pertains and as such will not be described further herein. For a given time t, the above calculation results into t/N sets of K spectral values of signal W 470, each spectral value being associated to a respective frequency band from a set of K frequency bands. Mathematically, this can be expressed as follows:

Equation 36

$$W_{SPECTRA} = FFT(w(t)) = \begin{bmatrix} W_{SPECTRA}(0) \\ W_{SPECTRA}(1) \\ W_{SPECTRA}(2) \\ \dots \\ W_{SPECTRA}(K-2) \\ W_{SPECTRA}(K-1) \end{bmatrix}_0 \begin{bmatrix} W_{SPECTRA}(0) \\ W_{SPECTRA}(1) \\ W_{SPECTRA}(2) \\ \dots \\ W_{SPECTRA}(K-2) \\ W_{SPECTRA}(K-1) \end{bmatrix}_1 \dots \begin{bmatrix} W_{SPECTRA}(0) \\ W_{SPECTRA}(1) \\ W_{SPECTRA}(2) \\ \dots \\ W_{SPECTRA}(K-2) \\ W_{SPECTRA}(K-1) \end{bmatrix}_{\frac{t}{N}-1}$$

where

$$W_{j,SPECTRA} = \begin{bmatrix} W_{SPECTRA}(0) \\ W_{SPECTRA}(1) \\ W_{SPECTRA}(2) \\ \dots \\ W_{SPECTRA}(K-2) \\ W_{SPECTRA}(K-1) \end{bmatrix}_j$$

where $W_{SPECTRA}$ 412 is a data structure of t/N vectors each of size K, each vector being indicative of a spectral representation of N samples signal W 470 and $W_{SPECTRA}(j)$ is the spectral value of signal W 470 associated to frequency band j. $W_{SPECTRA}$ 412 is released by spectral calculator 404.

Methods other than the FFT for dividing a signal into a set of frequency bands may be used by the spectral calculators 404, 406, such as for example, a cosine transform and other similar transforms. Although spectral calculator 406 and spectral calculator 404 are depicted as separate components in figure 4, it will be readily appreciated that they may be embodied in a same physical device and may share functional components without detracting from the spirit of the invention.

The per-band standard deviation computation unit 408 receives W_{SPECTRA} 412 and Z_{SPECTRA} 410 and processes each frequency band to generate an error characterization estimate $\text{Herror}[j]$ for each band j , for $j=0 \dots K-1$. In a specific implementation, $\text{Herror}[j]$ is the standard deviation of error function for frequency band j . The per-band standard deviation computation unit 808 also generates a per-band energy estimate for signal Z 102, identified as Z_{energy} 250 in figure 8.

Figure 5 shows a conceptual block diagram of the per-band standard deviation computation unit 408. As depicted, the per-band standard deviation computation unit 408 includes a set of K parallel computation units 500 where each unit 500 is operative to compute the standard deviation of the error function for a respective frequency band. If the frequency bands are narrow, the signals Z 102 and W 470 can be considered “white” within a frequency band, thereby allowing the following computation to be used:

Equation 37

for $j=0 \dots K-1$.

$$H_{\text{error}}[j] = \frac{\sum_{i=0}^{\frac{t}{N}-1} ((W_{i,\text{SPECTRA}}[j] \times Z_{i,\text{SPECTRA}}[j])^2)^{\frac{1}{2}}}{\left[\left(\sum_{i=0}^{\frac{t}{N}-1} Z_{i,\text{SPECTRA}}[j] \times Z_{i,\text{SPECTRA}}[j] \right)^2 - \left(\sum_{i=0}^{\frac{t}{N}-1} (Z_{i,\text{SPECTRA}}[j] \times Z_{i,\text{SPECTRA}}[j])^2 \right) \right]^{\frac{1}{2}}}$$

where Herror[j] is the error characterization data element for frequency band j and Herror
 208 is a set of K error characterization data elements. Each unit 500 also releases a data
 5 element Zenergy[j] indicative of the energy of signal Z 102 in frequency band j and
 Zenergy 250 is a set of K energy data elements. Zenergy is computed as follows:

Equation 38

$$Z_{energy}[j] = \sum_{i=0}^{t-1} (Z_{i,SPECTRA}[j] \times Z_{i,SPECTRA}[j]) \quad \text{for } j = 0 \dots K-1$$

10

The skilled person in the art will readily appreciate that the implementation
 depicted in figure 4 is for the purpose of example only as many other implementations are
 possible.

15

Although the above described specific examples of implementations show the
 computations in the frequency domain of the auto-correlation of signal Z 102 and the
 cross-correlation of signals Z 102 and W 470, it is to be understood that the equivalent of
 either of these computations may be effected in the time domain without detracting from
 the spirit of the invention. For example, the auto-correlation and cross-correlation
 20 computations may be effected in the time domain while the computation of the standard
 deviation is effected in the frequency domain.

Other examples of implementation of the error characterization unit 202 are
 described in United States Patent Application entitled, "Method and Apparatus for
 25 Providing an Error Characterization Estimate of an Impulse Response Derived using
 Least Squares", filed on the same date as the instant application by Awad T. et al. The
 content of the above referenced document is hereby incorporated by reference.

In the alternative non-limiting example where Hnew 206 comprises a plurality of
 30 sets of filter coefficients, Hnew[i], for i=0...FC-1, a set of error characterization data

elements is generated for each set of filter coefficients in the plurality of sets of filter coefficients using the above described procedure. In this alternative non-limiting example, *Error* is comprised of a plurality of sets of error characterization data elements, each set of error characterization data elements being associated to a respective set of filter coefficients. *Error* can be expressed as a vector of sets of error characterization data elements *Error*[*i*], for *i*=0...*FC*-1, where *FC* is the number of sets of filter coefficients in *Hnew* 206.

Error 208 and *Zenergy* 250 are released by the error characterization unit 202 and provided to the coefficient set selection unit 204 depicted in greater detail in figure 6.

Active Coefficient Set Memory Unit 240

The active coefficient set memory unit 240 stores a set of filter coefficients most recently selected by the coefficient set selection unit 204, herein referred to as *Hbest*, along with its associated set of error characterization data elements, herein referred to as *Hbest_error*. The *Hbest_error* includes *K* data elements, each data element being associated to a respective frequency band. The active coefficient set memory unit 240 stores the set of filter coefficients *Hbest* currently being released to the adaptive filter 110. *Hbest* is associated to an ordering data element, which will be referred to as *Hbest_age*, indicating the time of receipt of the first sample of signal *Z* 102 that was included in the computation of *Hbest*. At the time *Hbest* was generated, the context update module 300 (figure 3) included data derived from a first subset of samples of signal *Z* 102 and a first subset of samples of signal *X* 104. This may be the case when *Hnew* 206 includes one set of filter coefficients or when *Hnew* 206 comprises a plurality of sets of filter coefficients, *Hnew*[*i*], for *i*=0...*FC*-1.

In an alternative non-limiting example where *Hnew* 206 comprises a plurality of sets of filter coefficients, *Hnew*[*i*], for *i*=0...*FC*-1, the active coefficient set memory unit 240 *Hbest* comprises a plurality of sets of filter coefficients, *Hbest*[*i*], for *i*=0...*FC*-1, most recently selected by the coefficient set selection unit 204. *Hbest_error* includes a plurality of sets of error characterization data elements, *Hbest_error*[*i*], for *i*=0...*FC*-1,

where each set of error characterization data elements is associated to a corresponding set of filter coefficient in $Hbest[i]$, for $i=0...FC-1$. In this alternative non-limiting example, one of the plurality of sets of $Hbest$ in the active coefficient set memory unit 240 is currently being released to the adaptive filter 110. Each set of filter coefficients in $Hbest$ is also associated to an ordering data element, $Hbest_age$, indicating the time of receipt of the first sample of signal Z 102 that was included in the computation of corresponding set of filter coefficients. $Hbest_age$ comprises a plurality of ordering data elements, $Hbest_age[i]$, for $i=0...FC-1$.

Since $Hnew$ 206 was generated after $Hbest$, at least some samples of signals Z 102 and X 104 in the first subsets from which $Hnew$ 206 was generated were received by the filter adaptation unit 100 subsequently to all the samples in corresponding first subsets from which $Hbest$ was generated.

Coefficient Set Selection Unit 204

The coefficient set selection unit 204 is operatively coupled to the active coefficient set memory unit 240 to receive $Hbest$ along with its associated set of error characterization data elements, $Hbest_error$ and its corresponding ordering data element $Hbest_age$. The coefficient set selection unit 204 also receives the set of filter coefficients $Hnew$ 206 generated by the coefficient generation unit 200 as well as the associated ordering data element $Hnew_age$ and the associated set of error characterization data elements $Herror$ 208 generated by the error characterization unit 202. The coefficient set selection unit 204 compares $Hnew$ 206 to $Hbest$ as well as the set of error characterization data elements associated to $Hnew$ 206 and the set of error characterization data elements associated to $Hbest$ in order to select the set of filter coefficients H 116 to be released. The comparison criteria may be based on various criteria designed to select a set of filter coefficients that best describes system 150 (figure 1).

The coefficient set selection unit 204 will now be described in greater detail with reference to Figure 6, which depicts the coefficient set selection unit 204 in greater detail.

The coefficient set selection unit 204 includes memory unit 602, a detection of a change unit 604, an improvement detection unit 606, an existence detection unit 612 and a decision module 610.

- 5 The memory unit 602 receives and stores H_{new} 206, its associated set of error characterization data elements H_{error} 208 and its associated ordering data element H_{new_age} .

Detection of a change unit 604

10

The detection of a change unit 604 attempts to determine the likelihood that differences between a first set of filter coefficients H_1 and a second set of filter coefficients H_2 can be attributed to general statistical variability in the impulse response of a same state of system 150.

15

In a non-limiting example, the detection of a change unit 604 receives from the decision module 610 one set of filter coefficients H_1 associated to H_{best} and one set of filter coefficients H_2 associated to H_{new} , 206 along with their corresponding sets of error characterization data elements and ordering data elements.

20

- In a specific implementation, the ordering data element associated to H_2 , namely H_{2_age} , and the ordering data element associated to H_1 , namely H_{1_age} , are first compared to determine whether H_2 is associated to an ordering data element indicating a later time of receipt than the ordering data element associated to H_1 . More specifically, if
- 25 H_2 is associated to an ordering data element indicating an earlier time of receipt than the ordering data element associated to H_1 , the detection of a change unit 604 releases a signal indicating that no change has been detected to the decision module 610. If H_2 is associated to an ordering data element indicating a later time of receipt than the ordering data element associated to H_1 , the detection of a change unit proceeds with the processing
- 30 described below. The above test allows reducing the likelihood of older data causing the detection of a change in the state of the system where a set of filter coefficients based on newer data has already been obtained.

In a non-limiting example, H_2 and H_1 are divided into a set of K frequency bands. For each frequency band, the amplitudes of H_2 and H_1 are compared. More specifically, for each frequency band in the set of frequency bands, a first distance measure data element indicative of the difference in the spectral domain between the amplitudes of H_2 and H_1 is computed. Mathematically, this can be expressed as follows:

Equation 39

$$\text{DISTANCE_AMPLITUDE}[j] = \text{Distance}(H_1[j], H_2[j]) \quad \text{for } j = 0 \dots K - 1$$

10

where $\text{Distance}(x, y)$ is a distance computation between two complex numbers, $H_1[j]$ and $H_2[j]$ are the components in frequency band j of H_1 and H_2 respectively, and $\text{DISTANCE_AMPLITUDE}$ is a set of K first distance measure data elements.

15

The sets of error characterization data elements H_1_error and H_2_error are also processed. In this specific implementation, each error characterization data element includes a standard deviation data element. More specifically, for each frequency band in the set of frequency bands, a second distance measures data element indicative of the sum of the standard deviation data elements of H_1_error and H_2_error is generated.

20 Mathematically, this can be expressed as follows:

Equation 40

$$\text{ERROR_RANGE}[j] = H_1_error[j] + H_2_error[j] \quad \text{for } j = 0 \dots K - 1$$

25

Alternatively, since H_1 was generated on the basis of older samples of signals Z 102 and X 104 and H_1_error is static and is not affected by newer samples, H_1_error is not necessarily indicative of the performance of the set of filter coefficients H_1 on the current samples of Z 102. Consequently, in this alternative implementation, H_1_error can be removed from equation 40 indicating that no variation can be accepted on H_1 .

30 Mathematically, this can be expressed as follows:

Equation 41

$$ERROR_RANGE[j] = H_2error[j] \quad \text{for } j = 0 \dots K - 1$$

Following this, for each frequency band in the set of frequency bands, the ratio of
 5 the difference between the amplitudes of H_2 and H_1 and the sum of the standard deviation
 of the error function is calculated. Mathematically, this can be expressed as follows:

Equation 42

$$RATIO[j] = \frac{DISTANCE_AMPLITUDE[j]}{ERROR_RANGE[j]} \quad \text{for } j = 0 \dots K - 1$$

10 If the ratio for a given frequency band is greater than 1, then it can be concluded that it is
 likely that H_2 and H_1 are inconsistent for the frequency band being considered. For the
 set of K ratio, statistical methods can be used to determine where the ratios are located.

In a first non-limiting implementation, the set of ratios is attributed to a Gaussian
 15 distribution and well-known statistical tests are performed to determine the likelihood that
 the observed distribution of ratios are indicative of a distribution of ratios greater than 1
 with a given confidence level X%. If the statistical tests indicate that the observed
 distribution of ratios are indicative of a distribution of ratios greater than 1 with a given
 confidence level X%, then the detection of a change unit 604 releases a signal to the
 20 decision module 610 indicating that a change has been detected. Otherwise, the detection
 of a change unit 604 releases a signal to the decision module 610 indicating that no
 change has been detected.

In a second non-limiting example, a sum of the Zenergy 250 values for each
 25 frequency bands where the ratio is greater than a certain ratio threshold is computed and
 compared against a threshold. The sum is then compared to a portion of total energy of
 signal Z to determine whether a change in the state of system 150 has likely occurred.
 Mathematically this can be expressed as follows:

Equation 43

$$\begin{aligned}
 &RATIO_SUM = 0 \\
 &\text{for } j = 0 \dots K-1 \\
 &\quad \text{If } RATIO[j] > \alpha \\
 &\quad \quad RATIO_SUM = RATIO_SUM + Zenergy[j]
 \end{aligned}$$

Equation 44

$$TOTAL_Zenergy = \sum_{j=0}^{K-1} Zenergy[j]$$

5

where α is a certain ratio threshold. The ratio threshold α may be selected on the basis of trial and error or on the basis of heuristic performance measurements. In non-limiting implementation, α is set to about "1" however other values are possible without detracting from the spirit of the invention.

10

If $RATIO_SUM$ is greater than a certain percentage of the $Total_Zenergy$, then the detection of a change unit 604 releases a signal to the decision module 610 indicating that a change has been detected. Otherwise, the detection of a change unit 604 releases a signal to the decision module 610 indicating that no change has been detected. The

15 certain percentage is established by trial and error based on the percentage that gives the best results.

Improvement detection unit 606

20

The improvement detection unit 606 selects the set of filter coefficients that minimizes the error over all the frequency bands.

In a non-limiting example, the improvement detection unit 606 receives from the decision module 610 one set of filter coefficients H_1 associated to H_{best} and one set of

25 filter coefficients H_2 associated to H_{new} 206, along with their corresponding sets of error characterization data elements, H_1_error and H_2_error respectively.

More generally, the improvement detection unit 606 selects the set of filter coefficients that minimize a weighted sum of the error characterization data elements over all the frequency bands. Mathematically, this second example may be expressed as follows:

5

Equation 45

If $Weighted(H_2error) \leq Weighted(H_1error)$
 then change
 else no change

where

$$Weighted(H_2error) = \sum_{j=0}^{K-1} w_j H_2error[j]^2$$

$$Weighted(H_1error) = \sum_{j=0}^{K-1} w_j H_1error[j]^2$$

10 where w_j is a weight associated to frequency band j . In a specific example of implementation, the weight w_j is the energy of signal Z for frequency band j , namely $Zenergy[j]$. In other words, if the weighted sum of the error characteristic data elements of the new set of filter coefficients H_2 is less than or equal to the weighted sum of the error characteristic data elements of the old set of filter coefficients H_1 , then an
 15 improvement in the error characteristic is detected and the improvement detection unit 606 releases a signal to the decision module 610 indicating that an improvement has been detected. Otherwise, the improvement detection unit 606 releases a signal to the decision module 610 indicating that no improvement has been detected.

20 ***Existence detection unit 612***

The existence detection unit 612 determines whether a selected set of filter coefficients is indicative of a substantially null transfer function.

In a non-limiting example, the maximum amplitude of the frequency response described by a selected set of filter coefficients, say *Selected_H*, in each frequency bands is compared to a reference value, where the reference value is indicative of an almost null transfer function. In a non-limiting example, the maximum value of the amplitude of *Selected_H* in a given frequency band *j* is *Selected_H[j]* added to *Selected_Herror[j]*. Mathematically, this can be expressed as follows:

Equation 46

$$DISTANCE[j] = \text{Magnitude}(\text{Selected_H}[j]) + \text{Selected_Herror}[j] \quad \text{for } j = 0 \dots K-1$$

10

where *Magnitude(x)* is a magnitude computation for a complex number, *Selected_H[j]* and *Selected_Herror[j]* are the components in frequency band *j* of *Selected_H* and *Selected_Herror* respectively, and *DISTANCE* is a set of *K* first distance measure data elements.

15

Following this, for each frequency band in the set of frequency bands, the ratio of the maximum amplitudes of *Selected_H* over the reference signal is calculated. Mathematically, this can be expressed as follows:

Equation 47

$$RATIO[j] = \frac{DISTANCE[j]}{REFERENCE[j]} \quad \text{for } j = 0 \dots K-1$$

20

where *REFERENCE[j]* is the amplitude of the reference impulse response in frequency band *j*. The “almost null” transfer function depends on the application in which the filter adaptation unit 100 is used. In a non-limiting example, the adaptive system 170 is used in an echo canceling system and the almost null transfer function is a response having a gain of about -35dB or less. For the set of *K* ratio, statistical methods can be used to determine where the ratios are located. In a first non-limiting implementation, the set of ratios is attributed to a Gaussian distribution and well-known statistical tests are performed to determine whether the distribution of the ratios is low enough that it cannot

25

30

be explained by normal statistical variations. Various well-known statistical methods may be used here without detracting from the spirit of the invention and as such will not be described further. If the statistical tests indicate that the distribution of the ratios is low enough that it cannot be explained by normal statistical variations, then the existence
5 detection unit 612 releases a signal to the decision module 610 indicating that a substantially null transfer function has been detected. Otherwise, the existence detection unit 612 releases a signal to the decision module 610 indicating that no substantially null transfer function has been detected.

10 ***Decision module 610***

The decision module 610 receives the results from the detection of a change unit 604 and the improvement detection unit 606 and makes use of these results to select a set of filter coefficients. The functioning of the filter adaptation unit 202 will be described
15 with regard to the flow diagram of figure 7a.

For the purpose of simplicity, a first example of implementation will first be described where *Hbest* includes one set of filter coefficients and *Hnew* 206 includes either one of a single set of filter coefficients or a plurality of sets of filter coefficients. The
20 manner in which this detection of a change unit 604 may be extended to the case where *Hnew* 206 and *Hbest* include a plurality of sets of filter coefficients will be described later on in the specification.

In the case where *Hnew* 206 includes a plurality of sets of filter coefficients,
25 decision module 610 starts at step 703 where one set of filter coefficients is selected from the plurality of sets of *Hnew*[i], for $i=0 \dots FC-1$. The decision module 610 selects a set of filter coefficients from the plurality of sets of *Hnew*[i] at least in part on the basis of energy values associated to respective sets of filter coefficients. The set of filter coefficients that provides the strongest impulse response is selected. In a non-limiting
30 example, the energy values associated to respective sets of filter coefficients are conditioned at least in part on the basis of energy values associated to signal Z 102. Mathematically, a specific non-limiting implementation can be expressed as follows:

Equation 48

$$Selected_Hnew = \underset{i=0}{MAX}^{FC-1} \left[\sum_{j=0}^{K-1} (Hnew[i][j])^2 * Zenergy[j] \right]$$

- 5 where $Hnew[i][j]$ is the component of $Hnew$ 206 in the i th segment for the j th frequency band. Alternatively, the associated sets of error characterization data elements $Error[i]$, $i=0...FC-1$, may be used to condition the results of equation 48. Mathematically, a specific non-limiting implementation can be expressed as follows:

Equation 49

$$Selected_Hnew = \underset{i=0}{MAX}^{FC-1} \left[\sum_{j=0}^{K-1} (Hnew[i][j] - \beta Error[i][j])^2 * Zenergy[j] \right]$$

- 10 where $Error[i][j]$ is the component of $Error$ in the i th segment for the j th frequency band and β is a scale factor ≤ 1 . This step 703 will result in one selected set of filter coefficients, namely $Selected_Hnew$, along with its associated error characterization data element, $Selected_Error$, and ordering data element, $Selected_Hnew_age$. In the equation above, when $Hnew[i][j] - \beta Error[i][j]$ is less than 0, the subtraction is set to 0. For step 703, the decision module proceeds to step 701.

- 20 In the case where $Hnew$ 206 includes one set of filter coefficients, $Selected_New$ = $Hnew$ and the decision module 610 proceeds at step 700. At step 700, the decision module 610 queries the detection of a change unit 604 to determine whether a change between $Selected_New$ and $Hbest$ can be detected. If the signal returned by the detection of a change unit 604 indicates that a change has been detected, the decision module proceeds to step 702 and replaces the old set of filter coefficients $Hbest$, $Hbest_error$ and $Hbest_age$ in memory unit 240 by the new set of filter coefficients $Selected_New$ 206, $Selected_Error$ and $Selected_Hnew_age$. Following this, at step 704, the decision module 610 queries the existence detection unit 612 to determine whether a substantially

null transfer function has been detected. If the signal returned by the existence detection unit 612 indicates that a substantially null transfer function has been detected, the decision module 610 proceeds to step 706 where a “0” transfer function is stored in memory unit 240. If at step 700 the signal returned by the detection of a change unit 604
 5 indicates that no change has been detected, the decision module proceeds to step 710, where the decision module 610 queries the improvement detection unit 606 to determine whether an improvement between Hbest and Selected_H is present. If the signal returned by the improvement detection unit 606 indicates that an improvement has been detected, the decision module 610 proceeds to step 702 and replaces the old set of filter coefficients
 10 Hbest, Hbest_error and Hbest_age in memory unit 240 by the new set of filter coefficients *Selected_New* 206, *Selected_Herror* and *Selected_Hnew_age*. Following this, at step 704, the decision module 610 queries the existence detection unit 612 to determine whether the current Hbest is indicative of a substantially null transfer function. The decision module 610 considers the signal received from the existence detection unit
 15 612. If the signal indicates that a substantially null transfer function has been detected, the decision module 610 proceeds to step 706 where a “0” transfer function is stored in memory unit 240. If at step 710, the signal indicates that a “0” transfer function has not been detected, no change made to the transfer function is stored in memory unit 240.

20 The set of filter coefficients in the active coefficient set memory unit 240 is released in a format suitable for use by filter 110.

In the alternative non-limiting example where *Hnew* 206 and Hbest both comprise a plurality of sets of filter coefficients, *Hnew*[i] and *Hbest*[i], for $i=0 \dots FC-1$, the above
 25 described process is effected for each set of filter coefficients in *Hnew* 206. Referring to figure 7b for each set of filter coefficients in *Hnew* 206, the following steps are performed:

At step 750, the decision module 610 queries the detection of a change unit 604 to
 30 determine whether a change between *Hnew*[i] and *Hbest*[i] can be detected. If the signal returned by the decision module 610 indicates that a change has been detected, the decision module proceeds to step 752 and replaces the old set of filter coefficients

Hbest[*i*], *Hbest_error*[*i*] and *Hbest_age*[*i*] in memory unit 240 by the new set of filter coefficients *Hnew*[*i*], *Herror*[*i*] and *Hnew_age*[*i*]. If at step 750, the signal returned by the decision module 610 indicates that no change has been detected, the decision module proceeds to step 754. At step 754, the decision module queries the improvement detection unit 606 to determine whether an improvement between *Hbest*[*i*] and *Hnew*[*i*] is present. If the signal returned by the improvement detection unit 606 indicates that an improvement has been detected, the decision module 610 proceeds to step 752 and replaces the old set of filter coefficients *Hbest*[*i*], *Hbest_error*[*i*] and *Hbest_age*[*i*] in memory unit 240 by the new set of filter coefficients *Hnew*[*i*], *Herror*[*i*] and *Hnew_age*[*i*]. Following step 752, or if the test at the signal returned by the improvement detection unit 606 indicated that no improvement was detected, the decision module 610 determines whether all the sets of filter coefficients *Hnew*[*i*], for *i*=0...*FC*-1 have been processed. If any set of filter coefficients *Hnew*[*i*] remains to be processed, condition 756 is answered in the negative and the decision module 610 returns to step 750. Once all the sets of filter coefficients have been processed, the memory unit 240 will comprise updated sets of *Hbest*[*i*], *Hbest_error*[*i*] and *Hbest_age*[*i*], for *i*=0...*FC*-1.

Once all the sets of filter coefficients have been processed, condition 756 is answered in the affirmative and the decision module 610 proceeds to step 758 where one set of filter coefficients is selected from the plurality of sets of *Hbest*[*i*], for *i*=0...*FC*-1 in memory unit 240. The decision module 610 selects a set of filter coefficients from the plurality of sets of *Hbest*[*i*] at least in part on the basis of energy values associated to respective sets of filter coefficients. The set of filter coefficients, which provides the strongest impulse response, is selected. In a non-limiting example, the energy values associated to respective sets of filter coefficients are conditioned, at least in part, on the basis of energy values associated to signal Z 102. Mathematically, a specific non-limiting implementation can be expressed as follows:

Equation 50

$$Selected_H = \underset{i=0}{MAX}^{FC-1} \left[\sum_{j=0}^{K-1} (Hbest[i][j])^2 * Zenergy[j] \right]$$

where $Hbest[i][j]$ is the component of $Hbest$ in the i th segment for the j th frequency band. Alternatively, the associated sets of error characterization data elements $Hbest_error[i]$, $i=0...FC-1$, may be used to condition the results of equation 50. Mathematically, a specific non-limiting implementation can be expressed as follows:

Equation 51

$$Selected_H = \underset{i=0}{MAX}^{FC-1} \left[\sum_{j=0}^{K-1} (Hbest[i][j] - Hbest_error[i][j])^2 * Zenergy[j] \right]$$

where $Hbest_error[i][j]$ is the component of $Hbest_error$ in the i th segment for the j th frequency band and β is a scale factor ≤ 1 . This step 758 will result in one selected set of filter coefficients, namely $Selected_H$, along with its associated error characterization data element, $Selected_Herror$, and ordering data element, $Selected_Hnew_age$. In the equation above, when $Hbest[i][j] - \beta Hbest_error[i][j]$ is less than 0, the subtraction is set to 0.

Following this, at step 760, the decision module 610 considers queries from the existence detection unit 612 to determine whether $Selected_H$ has a substantially null transfer function. The decision module 610 considers the signal returned by the existence detection unit 612 in response to the query. If the signal indicates that a substantially null transfer function has been detected, the decision module 610 proceeds to step 762 where a "0" transfer function is stored in memory unit 240. If at step 762, the signal indicates that a no substantially null transfer function has been detected, no change is made to the transfer function stored in memory unit 240.

The set of filter coefficients in the active coefficient set memory unit 240 is released as H 116 in a format suitable for use by filter 110.

The above-described process for producing a set of filter coefficients can be implemented on a general purpose digital computer, of the type depicted in figure 8,

including a processing unit 802 and a memory 804 connected by a communication bus. The memory includes data 808 and program instructions 806. The processing unit 802 is adapted to process the data 808 and the program instructions 806 in order to implement the functional blocks described in the specification and depicted in the drawings. The
5 digital computer 800 may also comprise an I/O interface 810 for receiving or sending data elements to external devices. For example, the I/O interface may be used for receiving the first signal Z 102 and the second signal X 104.

Alternatively, the above-described process for producing a set of filter coefficients
10 can be implemented on a dedicated hardware platform where electrical/optical components implement the functional blocks described in the specification and depicted in the drawings. Specific implementations may be realized using ICs, ASICs, DSPs, FPGA, or other suitable hardware platform. It will be readily appreciated that the hardware platform is not a limiting component of the invention.

15 Although the present invention has been described in considerable detail with reference to certain preferred embodiments thereof, variations and refinements are possible without departing from the spirit of the invention. Therefore, the scope of the invention should be limited only by the appended claims and their equivalents.

20